DOCUMENT RESUME

ED 401 515 CG 027 440

AUTHOR Curran, Lisa; And Others
TITLE Assessment of American Indian Children as Measured by the SON-R and WISC-III.
PUB DATE 96
NOTE 17p.; Paper presented at the Annual Meeting of the National Association of School Psychologists (28th, Atlanta, GA, March 12-16, 1996).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *American Indians; Children; Cultural Context; *Culture Fair Tests; *Intelligence Differences; *Intelligence Tests; Test Validity
IDENTIFIERS Snijders Oomen Nonverbal Test of Intelligence Rev; Wechsler Intelligence Scale for Children III

ABSTRACT

A major criticism of standardized intelligence tests is their improper use in measuring the intellectual competence of culturally diverse children. Factors which complicate the issue are the definition of intelligence, content bias in intelligence tests, and the interpretation of test scores between white middle class children and children of other groups. This study compared the performance of American Indian children on two different intelligence tests: (1) the Wechsler Intelligence Scale for Children-Third Edition (WISC-III); and (2) the Snijders-Oomen Nonverbal Test of Intelligence - Revised (SON-R). Studies suggest that American Indian children performed poorly on the Verbal Scale of the WISC-III which subsequently lowered their Full Scale IQ scores. Since the SON-R is designed to de-emphasize the verbal component of IQ testing, it was hypothesized that the SON-R would more accurately assess American Indian children's intelligence than the WISC-III. Results indicated no significant difference between performance on the two tests, failing to confirm the hypothesis. Contains 12 references. (RB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 401 515

ASSESSMENT OF AMERICAN INDIAN CHILDREN AS MEASURED BY THE SON-R AND WISC-III (PA204)

Lisa Curran
Diane Elkerton
Mark Steinberg

University of Wisconsin-River Falls
Office of School Psychology
AMES Education Building
River Falls, WI 54022

Paper Presentation at the
National Association of School Psychologists
28th Annual National Convention

Hyatt Regency
Atlanta, Georgia

March 12-16, 1996

We are especially grateful to Dr. Peter Tellegen at the University of Groningen, The Netherlands, for providing his own personal SON-R (5 1/2-17) test kit; to Dr. Douglas K. Smith, Director of School Psychology at the University of Wisconsin - River Falls for his invaluable support and assistance; and to the children for their participation in the study and to the staff at the Division of Indian Work of Greater Minneapolis Council of Churches in Minneapolis, Minnesota, for their assistance.

BEST COPY AVAILABLE

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
L. Curran
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CG027440

Abstract

This study compared the performance of American Indian children on the WISC-III and SON-R (5 1/2 - 17) intelligence tests. Studies suggest that American Indian children performed poorly on the Verbal Scale of the WISC-III which subsequently lowered their Full Scale IQ scores. Since the SON-R is designed to de-emphasize the verbal component of IQ testing, it was hypothesized that the SON-R would more accurately assess American Indian children's intelligence than the WISC-III.

The WISC-III mean FSIQ was 95.5 (sd = 9.6) and the mean SIQ for the SON-R was 97.4 (sd = 12.7), showing no significant difference. Correlations after corrections for restriction of range were: FSIQ and SIQ = .68.

A major criticism of standardized intelligence tests is their improper use in measuring the intellectual competence of culturally diverse children (Armour-Thomas, 1992). A complicating factor in this issue is that what has been designated and defined as intelligence in standardized tests may not be the same for all cultures. Variability of life experiences between and within culturally diverse groups occurs frequently. This includes access to educational resources (the schools that are available) and development of cognitive competencies other than those measured on standardized tests. For example, Gardner (1983) proposed that there are multiple intelligences, composed of seven abilities, talents, and mental skills: musical, bodily-kinesthetic, logical-mathematical, linguistic, spatial, interpersonal, and intrapersonal.

Another factor to consider is content bias. Test items from the dominant culture may measure concepts or contain text that is not available or used in other cultures. The more familiar and culturally appropriate an item is, the more likely the student will remember that item, have been socialized with that item, and react to it as expected. Also, language differences which can include "sociolinguistic patterns" that exist in different cultures, may not exist in the protocol of the test (Armour-Thomas, 1992). While test items may be assumed to be universally common to subjects' experience, culturally different children may be encountering language that is not common to their day to day activities. The linguistic difference, while perhaps not totally foreign, may be sufficiently different to influence test performance.

An additional issue of concern linked with intellectual assessment practices is the interpretation of differences in the test scores between non-white, low-income children and white, middle-class children. Critics question whether standardized tests of intelligence accurately reflect the cognitive or intellectual potential of the culturally diverse child. While many studies of intelligence assessment bias have been conducted with groups of African-American and Spanish children, there have been relatively few studies with American Indian children.

Commonly used norm-referenced measures are currently being used to measure IQ and make placement decisions about American Indian children. McShane and Plas (1984), found that the Wechsler Scales are the most commonly used tests of intelligence with American Indian populations. Test score differences between American Indian children and those on normed tests has been a research topic for the last 70 years. For example, Zarske and Moore (1982) studied the results of WISC-R performance of Navajo children. The children in this study were found to perform poorly on the Verbal Scale of the Wechsler Intelligence Scale for Children - Revised (WISC-R), and, as a result, the poor performance lowered their Full Scale IQ (FSIQ). According to the authors, the poor performance on the Conceptual scores was the direct result of cultural factors and English being a second language.

Naglieri & Yazzie (1983), examined the relationship between the WISC-R and the Peabody Picture Vocabulary Test-Revised (PPVT-R), a test that has been widely used as a measure of verbal comprehension. The subjects were Navajo students, the majority of whom spoke English as a second language. The children came from traditional Navajo families who had lived on the Navajo Reservation for many years and were low SES (only half of the subjects resided in dwellings that had electricity and running water). The Navajo children yielded a significantly lower mean score on the PPVT-R (mean = 61.1) than on the WISC-R (mean FSIQ = 87.4). Thus, the authors suggested that the PPVT-R should not be used as a measure of intelligence for American Indian children who speak English as a second language. They also recommended not using the Verbal Scale IQ (VIQ) of the WISC-R as a measure of intelligence because of the language skills influence.

Teeter, Moore, and Petersen (1982) examined WISC-R scores of American Indian (primarily Navajo) students referred for learning problems. Their study supported the following important conclusions:

> (a) The Performance Scale of the WISC-R can be used as the least biased measure of potential for non-LD Navajo children. (b) Navajo children, regardless of their specific learning deficits, score lower than the standardization group on verbal subtests. Consequently, the low verbal scores should be interpreted as reflecting divergent language, cultural, and experiential factors, rather than as a deficit in intellectual potential. However, it is possible that the Verbal IQ can be used as an index to determine the academic proficiency of Native American children in English-speaking classrooms, and specifically on reading comprehension tasks. Results also indicate that verbal tasks requiring attention and concentration are less difficult than those requiring a higher level competency in the English language. (c) Due to the influence of language, the Full Scale IQ yields a biased measure of intelligence and should not be reported for Navajo children as an overall index of intellectual functioning. (d) LD Navajo children with visual processing deficits demonstrate significantly lower performance IQs than those of their non-LD counterparts. Due to the specific perceptual disabilities negatively affecting performance tasks, and to the depressed English verbal abilities, the WISC-R is limited in its scope to obtain a measure of the real potential of these students. (p. 43-44).

American Indian children as a group score lower on the standard norm-referenced measures and the literature suggests that this is so because they tend to score lower on the Verbal parts of the test. They are not as familiar with the language used by the dominant culture.

More recently, several studies have compared the Wechsler Intelligence Scale for Children - Third Edition (WISC-III) with measures of nonverbal intelligence. At the First Annual South Padre Island Conference on Cognitive Assessment of Children and Youth in School and Clinical Settings on November 26-27, 1993, Tellegen, (1993), cited a study by Nieuwenhuys in 1991 comparing scores of 35 children from a university psychiatric clinic tested with the Snijders-Oomen Nonverbal Test of Intelligence - Revised (5 1/2-17) (SON-R), the WISC-R, and the Raven Progressive Matrices Test. Correlations between the SON-R Specific IQ (SIQ) and the WISC-R FSIQ and Performance IQ (PIQ) were .80. The correlation between the SON-R SIQ and WISC-R VIQ was .66. The SON-R mean SIQ score was 97.8 (sd = 15.8) and the WISC-R FSIQ was 95.6 (sd = 15.9).

Tellegen also cited a study by Jansen in 1991 (Tellegen, 1993). Jansen compared four verbal subtests of the WISC-R with the SON-R (5 1/2-17) on children diagnosed as having specific language impairment. Analysis of the results showed a large discrepancy between VIQ on the WISC-R (mean = 83.1; sd = 14.8) and nonverbal intelligence as measured with the SON-R (mean = 97.5; sd = 14.0). The difference was significant at the .01 level.

For fair assessment of ethnic minority children and children with language problems, non-verbal intelligence tests are generally recommended (Tellegen & Laros, 1993). The Snijders-Oomen tests of Non-verbal Intelligence have been used for this purpose since 1943. The latest revision of the SON-R (5 1/2-17) appears well suited for use in assessing children from minority populations.

The present study, therefore, was designed to investigate the performance of American Indian children on the WISC-III and the SON-R (5 1/2-17). Since the SON-R is designed to de-emphasize the verbal component of IQ testing, it is hypothesized that American Indian children will score higher on the SON-R than on the WISC-III.

Methods

Subjects

Twenty-eight American Indian school-aged children ranging in age from 6 years to 17 years participated in this study. Thirty subjects were randomly selected from those students participating in a community program through the Division of Indian Work in Minneapolis, Minnesota. Two subjects did not complete their participation, leaving a total of twenty-eight subjects. Subjects were stratified to form equal groups of males and females for age ranges 6-9 years, 10-13 years, and 14-17 years. Thirteen females and 15 males participated in this study. There were 13 subjects in the 6-9 year age range, 12 in the 10-13 year age range, and 3 subjects in the 14-16 year age range. Subjects came from various tribes including: Ojibwe, Sisseton Wahpeton Dakota,

Winnebago, Oglala Lakota, Kickapoo, and Sac & Fox. English was the primary language for all subjects.

Instrumentation

The WISC-III and the SON-R were used as assessment comparison measures. The WISC-III was standardized in the United States on 2200 children (Wechsler, 1991). There were 100 boys and 100 girls in each of 11 age groups, ranging from 6 to 16 years of age. The standardization sample was stratified on age, geographic region, parent education (used as an estimate of SES), and race/ethnicity. Children were classified for race/ethnicity membership as white, black, Hispanic, and Other (Native American, Eskimo, Aleut, Asian, and Pacific Islander). Geographic regions sampled were Northeast, North Central, South and West. For each age group, children were selected to as closely as possible match the proportions found in the 1988 U.S. Census data with respect to race/ethnicity, parent education, and geographic region.

Most of the white and black subjects in the standardization sample came from the North Central and South regions, the majority of the Hispanic and Other subjects came from the West and South (Wechsler, 1991). Race and ethnic proportions in the WISC-III standardization sample were 70.1% white, 15.4% black, 11.0% Hispanic, and 3.5% Other.

The three scales on the WISC-III (Verbal, Performance, and Full Scale) have internal consistency reliability coefficients of .89 or above for all the age ranges included in the standardization group (Wechsler, 1991). For the eleven age groups, the average internal consistency reliability coefficients were .96 for the Full Scale IQ, .91 for the Performance Scale IQ, and .95 for the Verbal Scale IQ.

Concurrent validity studies between the WISC-III and WISC-R are reported in the WISC-III manual (Wechsler, 1991). A sample of 206 children between ages 6 and 16 were administered the WISC-R and WISC-III in counter-balanced order. Correlations

were .81 for the Performance Scale IQ, .90 for the Verbal Scale IQ, and .89 for the Full Scale IQ.

The WISC-III manual provides data that supports the construct validity of the test (Wechsler, 1991). Factor analytic data cited in the manual indicates that the WISC-III adequately measures two factors that correspond to the Verbal and Performance Scales of the test. The data also suggest that the WISC-III is a fair measure of general intelligence.

The SON-R (5 1/2-17) was standardized in the Netherlands on 1350 subjects between the ages of 6 and 14 years (Laros & Tellegen, 1991). There were 150 subjects included in each of nine age groups. These groups were stratified by demographic variables, sex and educational type. Subjects with severe mental and physical handicaps were not included in the standardization sample.

The SON-R (5 1/2-17) reported reliability coefficients of .76 on average (after correction) for its individual subtests (Tellegen & Laros, 1993). The most reliable subtests were Mosaics, Patterns and Analogies. Total test score reliability was found to be .93. The SON-R has a mean of 100 and a standard deviation of 15 (same as the WISC-III).

Substantial data have been gathered to support the validity of this test (Tellegen and Laros, 1993). According to the authors, there is a strong relation between school achievement and intelligence and because of their relationship, school success is considered to be an indication of the validity of an IQ measure. A stepwise multiple regression analysis was used to examine the relationship between test performance and school achievement. The results indicate a significant difference in test scores between general education and special education students, and grade repetition was strongly related to the IQ scores.

The SON-R individual subtests are as follows (Laros & Tellegen, 1991):

- **Categories**: In this subtest the subject is shown three drawings and then is required to identify two drawings that have the same underlying concept (i.e., fruit) from five alternatives.
- **Mosaics**: In this subtest subjects copy mosaic patterns using nine red/white squares.
- **Hidden Pictures**: For this subtest subjects find hidden items in a drawing (i.e., kites).
- **Patterns**: In this subtest subjects are required to draw missing parts of a pattern.
- **Situations**: For this subtest subjects are shown a picture of a situation with missing parts and must choose the correct parts in order for the situation to make sense.
- **Analogies**: In this subtest subjects are required to discover analogies within sets of geometric figures (i.e., A:B = C:D).
- **Stories**: For this subtest subjects are shown picture cards and must correctly sequence them to form a logical story.

Procedure

The WISC-III and the SON-R (5 1/2-17) were individually administered to all subjects in counterbalanced order. The examiners were trained in the administration of both the WISC-III and the SON-R and followed all procedures specified in the test manuals. Testing began August 23, 1994 and was completed on September 7, 1994. The mean time interval between test administrations was 7.2 days with a range of 3 to 14 days.

Results/Discussion

FSIQ scores ranged from 75 to 110 with a mean score of 95.5 (sd = 9.6). SIQ scores range from 77 to 126 with a mean score of 97.4 (sd = 12.7). PIQ scores range from 81 to 121 with a mean score of 102.3 (sd =11.5). VIQ scores range from 70 to 111 with a mean score of 90.5 (sd = 11.5). In order to determine if the FSIQ and SIQ scores differed significantly from each other, T-tests for related samples were conducted with nonsignificant results. Pearson product moment correlations between SON-R SIQ and WISC-III IQ scores were calculated after corrections for restriction of range, and produced the following results: FSIQ and SIQ = .68 ($p<.001$), PIQ and SIQ = .52 ($p<.01$), and VIQ and SIQ = .51 ($p<.01$).

Pearson product moment correlations for individual subtest scores revealed some significant correlations between WISC-III and SON-R subtests. The subtests with the highest correlations were Arithmetic on the WISC-III and Patterns on the SON-R (.66, $p<.001$), followed by Block Design and Mosaics and Arithmetic and Situations (.60, $p<.001$). All other subtest correlations were less than .60. The complete list of correlations is presented in Table 1.

Insert Table 1 about here

The hypothesis that American Indian children would score higher on the SON-R (5 1/2-17) than the WISC-III was not confirmed. Scores on both tests were in the average range and not significantly different from each other. Based on these results, this suggests that the SON-R does not deemphasize the verbal component in IQ testing when compared to WISC-III FSIQ scores. Because correlations between FSIQ and SIQ were high, and the differences in mean scores was minimal, it appears the two tests are measuring intelligence in similar ways and are essentially equivelant to each other. The pattern of correlations suggest that the SIQ is measuring elements of verbal and nonverbal intelligence. The idea that because the test responses on the SON-R are not verbal in nature does not necessarily mean the reasoning used is not language based. The assumption that because the SON-R is labeled as a nonverbal measure may not necessarily be accurate and may not establish that the SON-R is a non-language measure.

Although there is a significant relationship between the global scales, there are few related subtests. Out of 84 possible subtest correlations, only 22 were significant (26%). Arithmetic correlates significanly with every SON-R subtest. Block Design correlates with 4 out of 7 subtests, while Vocabulary and Symbol Search each correlate with 3 out of 7 subtests. Digit Span and Coding only correlate with one subtest on the SON-R. Perhaps the SON-R subtests are measuring fluid intelligence that is not measured by the WISC-III. However, the majority of SON-R subtests do not correlate with some of the WISC-III subtests (i.e., Comprehension, Picture Arrangement, Picture Completion, and Similarities). Although the SON-R is described as a nonverbal instrument, almost half of the significant subtest correlations were with verbal subtests on the WISC-III. This indicates that much of the overlap in subtests is in the verbal area. At the same time, the small number of significant subtest correlations seems to suggest the SON-R is measuring aspects of intelligence that are different from the WISC-III.

Another factor influencing the results are subject characteristics. The subjects in the study may not have been typical of the American Indian population as they were children who live in an urban area where the population is largely comprised of individuals from the dominant culture. Their opportunities, both in and out of school, to interact with dominant culture children are much greater than American Indian children who reside in areas where the population's cultural background is largely unblended (ie. children who reside on a reservation). It is possible the reason the subjects scored more like dominant culture children is that they have had similar opportunities and education as their dominant culture neighbors.

There were no American Indian subjects participating in the standardization of the SON-R (5 1/2-17). The SON-R was standardized on children in the Netherlands, and therefore, this test may not be as valid when testing children in the United States. If the SON-R were restandardized on a stratified sample representative of the population of the United States, including American Indians, it may be a more valid measure when used with American children.

Results of this study should be interpreted with caution since the sample size was relatively small. Additional study samples could include a larger sample size, and be tribe-specific. Also, further studies could focus on American Indian children that reside on a reservation rather than in a large metropolitan area. Overall, the results of this study found no significant difference between subject scores on the SON-R (5 1/2-17) and WISC-III. Since the SON-R has the capability of being administered nonverbally (using only gestures for subtest directions), replication of this study utilizing SON-R nonverbal administration procedures may yield different results. All verbal components would then be eliminated during the administration procedures. These proceedures may be more effective in diminishing language effects.

Further studies comparing the SON-R to other measures of intelligence, such as the Differential Abilities Scale (DAS), Woodcock-Johnson - Revised (WJ-R), or the

Kaufman Adult and Adolescent Intelligence Test (KAIT), may yield different results. These three instruments are based on theories of intelligence including such factors as crystallized, fluid, and visual intelligence as described by Carroll, as well as the Horn/Cattell theory. These intelligence measures include subtests which may resemble SON-R subtests, such as Mosaics on the SON-R and Pattern Construction on the DAS or Analogies on the SON-R and Logical Steps on the KAIT, which may resemble SON-R subtests more closely than the WISC-III.

REFERENCES

Armour-Thomas, E. (1992). Intellectual assessment of children from culturally diverse backgrounds. School Psychology Review, 21, 552-565.

Flaugher, R. L. (1978). The many definitions of test bias. American Psychologist, 33, 671-679.

Gardner, H. (1983). Frames of mind: The theory of multiple intelligences. New York: Basic Books.

Jansen, B. (1991). A nonverbal alternative to the Wechsler Scales: The Snijders-Oomen Nonverbal Intelligence Test. In a Paper presented at the First Annual South Padre Island Conference on Cognitive Assessment of Children and Youth in School and Clinical Settings.

Laros, J.A., & Tellegen, P.J. (1991). Construction and validation of the SON-R 5½-17, the Snijders-Oomen non-verbal intelligence test. Groningen, The Netherlands: Wolters-Noordhoff.

McShane, D. A., & Plas, J. M. (1984). The cognitive functioning of American Indian children: Moving from the WISC to the WISC-R. School Psychology Review, 13, 61-73.

Naglieri, J. A., Yazzie, C. (1983). Comparison of the WISC-R and PPVT-R with Navajo children. Journal of Clinical Psychology, 39, 598-600.

Nieuwenhuys, M. (1991). A nonverbal alternative to the Wechsler Scales: The Snijders-Oomen Nonverbal Intelligence Test. In a Paper presented at the First Annual South Padre Island Conference on Cognitive Assessment of Children and Youth in School and Clinical Settings.

Teeter, A., Moore, C. L., & Petersen, J. D. (1982). WISC-R verbal and performance abilities of Native American students referred for school learning problems. Psychology in the Schools, 19, 39-44.

Tellegen, P. J., & Laros, J. A. (1993). The construction and validation of a nonverbal test of intelligence: The revision of the Snijders-Oomen tests. European Journal of Psychological Assessment, 9, 147-157.

Wechsler, D. (1991). Wechsler Intelligence Scale for Children - Third Edition (WISC-III), San Antonio, TX: The Psychological Corporation.

Zarske, J. A., & Moore, C. L. (1982). Recategorized WISC-R scores of learning disabled Navajo Indian children. Psychology in the Schools, 19, 156-159.

Table 1 WISC-III and SON-R (5 1/2-17) Subtest Correlations

Subtest	ANA	CAT	HID	MOS	PAT	SIT	STO
ARI	.55**	.59***	.54**	.52**	.66***	.60***	.45**
BD	.20	.33	.42*	.60***	.38*	.46*	.11
COM	-.14	.07	.08	.08	-.03	-.00	.19
CD	-.20	.07	.04	.13	.12	.13	-.39*
DS	.36	.34	.21	.33	.43*	.17	.04
INF	.18	.40*	.01	.26	.35	.17	.06
OA	.09	.36	.35	.44*	.23	.43*	-.02
PA	-.08	.16	.27	.29	.28	.30	.32
PC	.26	.20	.27	.29	.31	.33	.19
SIM	.17	.32	.09	.25	.24	.24	.27
SS	.34	.27	.57**	.46*	.51**	.37	.09
VOC	.27	.46*	.10	.41*	.37	.47*	.24

NOTE * p<.05
** p<.01
*** p<.001

Subtest Abbreviations

SON-R Subtests

ANA = Analogies
CAT = Categories
HID = Hidden Pictures
MOS = Mosaics
PAT = Patterns
SIT = Situations
STO = Stories

WISC-III Subtests

COM = Comprehension
CD = Coding
DS = Digit Span
INF = Information
OA = Object Assembly
PA = Picture Arrangement
PC = Picture Completion
SIM = Similarities
SS = Symbol Search
VOC = Vocabulary